PC Support Advisor

The Essential Resource for PC Support Professionals

Update 68

How To Speed Up Microsoft Access
The faster, the better 3
Mike Lewis provides 20 tips for speeding up Microsoft's database application for Windows.
File: A0618.1

Understanding Optical Storage Technology
WORM, CD-R and more. What's it all about? 7
Dave Stott launches an investigation into the workings of optical storage products.
File: H0812.1

Understanding Colour Notebook Technology
Active-matrix and Passive-matrix explained 15
Mike Lewis explains the different screen technologies used in colour notebook PCs.
File: H0813.1

How To Buy And Install CD-ROM Drives
SCSI or proprietary, the process is similar......19
Julian Moss offers some advice on safe and correct installation of CD-ROM drives.
File: H0211.1

Understanding The Clipper Chip
Stand aside DES and RSA......23
If Clinton gets his way, the only method of encrypting your data will be with a device that the US Government knows how to crack.
File: C0212.1

On This Month's Utility Disk

MegaWord - Document management tools for WinWord 6.
Comprehensive database of Seagate drive settings and jumper positions.
Troubleshooting guide for Delrina fax software.

"Problem solved" - says Robert Schifreen.

A few years ago, a clever marketing person invented the idea of the solution. The concept changed the way that computers and software were sold.

Nowadays, you don't buy a computer and a copy of Access. Instead, you buy a solution to your data management problem. (If the salesman is on form, it's a solution to a problem that you never actually had, but that's another story.)

In the vast majority of cases, computing solutions tend to be based on IBM PC compatible technology and Windows software. Whatever the problem you have, someone somewhere will sell you a solution based on PCs and Windows. Or, in a limited number of cases, Apple Macs.

Although a PC running Access is a satisfactory solution to the problem of data management, there are, of course, a number of other solutions. Few would deny that a Power Mac, based on the latest Power PC chips, is technically superior to a 486 or Pentium PC running Windows 3.1 or even NT. And a top-of-the-range DEC Alpha box, too, packs more power than any single-processor Intel-based PC. Especially when that Intel box is running MS-DOS, an operating system that's 13 years old and still can't do anything with a Pentium chip apart from running it in 8086 mode.

So why, if Power PC and DEC Alpha technology is better than Intel boxes, are we still buying Intel? Why do so many of the world's leading companies choose to run their business (or, indeed, risk their business) on old-fashioned technology? And why are sales of Alpha and Power PC machines so small in comparison?

Because, in so many cases, the Intel/Microsoft solution is just that. A solution to the problem. And once you've solved someone's problem, it's very very difficult to persuade them that it needs solving again, in a different way.

An IBM compatible PC running MS-DOS and Windows, or even OS/2 or NT, is not the best use of technology. But it's a solution, and one that works for millions of people.

If Apple, IBM and DEC think for a moment that they can, within the next decade or so, start to convert users away from established technology and onto "better" solutions, I think they're mistaken.

My guess is that the future belongs to Intel and Microsoft. However, I'd also suggest that a PC with a single processor will be as rare, five years from now, as a machine with 256 KB of RAM is today.

PCSA

PC Support *Advisor*

PC Support *Advisor* is a publication of **International Technology Publishing**.

ISSN: 1031-3966

Editor: Robert Schifreen

Editor in Chief: Les Bell

Publisher: Chris Carvan

Email:
hex@cix.compulink.co.uk
CompuServe: 100016,301

Copyright:
All material is *Copyright © 1994 International Technology Publishing*, and must not be reproduced in part or full, by any means, without the written permission of the publisher.

Liability:
The information, opinions, interpretations and directions herein are based on the best information available, but their completeness and accuracy cannot be guaranteed.

Responsibility:
Responsibility for the views and opinions expressed in this publication lies with the publisher, International Technology Publishing, and not with any agents of the publisher nor the editors or contributing editors.

International
Technology
Publishing Inc.
North American Subscriptions
Suite 200
555 De Haro Street
San Fransisco, CA94107
USA

Telephone: (415) 255 1295
Facsimile: (415) 255 8496

International
Technology
Publishing Inc.
European Subscriptions
13 Callcott Street
London
W8 7SU
United Kingdom

Telephone: (071) 724 9306
Facsimile: (071) 706 3296

International
Technology
Publishing Pty Ltd.
SE Asian Subscriptions
138 Cecil Street
Cecil Court #13-03
Singapore
0106

Telephone: (65) 227 4187
Facsimile: (65) 227 2885

International
Technology
Publishing Pty Ltd.
Australasian Subscriptions
Suite 511
185 Elizabeth Street
Sydney, NSW, 2000
Australia

Telephone: (02) 261 4683
Facsimile:(02) 261 4741

How To Speed Up Microsoft Access

Mike Lewis provides 20 tips for speeding up Microsoft's database application for Windows.

Microsoft Access boasts an impressive range of features, but slick performance is not one of them. If you support users of this popular Windows database, you will probably have heard many complaints of its sluggishness and lethargy. So here are a score of tips for speeding up this ponderous program.

1. Add memory

The one sure-fire method of improving Access's performance is to install more memory. Microsoft recommends at least 4 MB of RAM for Access 1.x, and 6 MB for Access 2.0. But those figures should be regarded as the absolute minimum. For adequate performance, you really need 8 MB; 12 MB is better still.

2. Maximise your resources

If adding that much RAM is not possible, you should at least free up as much of your existing resources as you can. All the usual performance-tweaking tips apply here, so remove your wallpaper, set your desktop background to "none", and disable your screen saver. Keep your Program Manager groups to a minimum and cut down on the TrueType fonts.

You can also boost performance by creating a permanent (rather than temporary) swap file on the fastest available drive, and enabling 32-bit disk access (and 32-bit file access if you have Windows 3.11) where possible. Finally, minimise or exit any other applications that are running at the same time as Access.

3. Reduce desktop clutter

During the course of an Access session, the workspace quickly becomes littered with open tables, queries, forms and reports. Get into the habit of closing all open objects as soon as you have finished with them. Even a small form or a nearly empty table will consume around 100 KB of memory, leading to a general slowing of the system. Objects open in design mode consume even more of resources.

4. Live without wizards

Wizards are a great boon, especially in designing forms and reports. But they consume resources at a fearful rate, even when not in use. If you can manage without these handy helpers, pull the plug on them.

In Access 2.0, wizards are implemented as add-ins libraries. To uninstall them, open the File menu and select Add-ins, then Add-in Manager. In the resulting dialogue (Figure 1), select the libraries that you wish to remove, and press the Uninstall button. The change takes effect the next time you start Access.

In Access 1.x, you remove wizards by deleting the "wizards.mda=ro" line from the [Libraries] section of MSACCESS.INI, which you'll find in your Windows directory.

5. Slim down the database

Like the desktop, the database can also become a dumping ground for old tables, forms, reports and queries. It's worth devoting a few minutes to reviewing the contents of the database. If you systematically delete the objects that you no longer need, you will end up with a smaller file, and this in turn will speed up processing.

Having removed unwanted objects, compact the file to get rid of space-wasting fragmentation. To do so, first close the database (and make sure that other users on the network close it too). Next, select Compact Database from the File menu. Select the database that you want to compact. Access will then prompt you for a name for the database to compact "into". You would normally specify the same name as the existing file, in which case the compacted version will replace the original. The rest of the process is automatic.

6. Use data-entry mode

When adding new records to a large table via a datasheet or form, switch to data-entry mode. You do this by selecting Data Entry from the Records menu. Access will hide the existing records, making it much faster to add new data. To return to normal, select Show All Records from the Records menu.

> *"You can boost performance by creating a permanent (rather than temporary) swap file on the fastest available drive, and enabling 32-bit disk access (and 32-bit file access if you have Windows 3.11)."*

7. Import rather than attach

Access provides two ways of getting at data that is held outside your database: importing and attaching. If you import a table, it is converted to Access's native format and stored in the database. If you attach it, the data remains in its original format and location; your database simply contains a pointer to it.

Working with imported tables is always faster than using attached data. To import a table, switch to the main database window, select Import from the File menu, then choose a data source (Figure 2). Select the file that you wish to import, and press the Import button. If the file is password-protected within its parent application, you will be prompted to enter the password at this point.

On completion of the import, the data will appear as a normal table within the database. You can go ahead and use it just like any other table. Keep in mind, though, that this is a copy of the data. Any changes you make to it will not affect the original. If that is not what you want, attaching might be a more appropriate option than importing.

8. Uses indexes carefully

Think carefully about which fields you want to index. Indexes can speed up sorts, searches and queries, but they tend to slow down data entry. You'll probably want to index a field if you run frequent queries on it, especially if it contains many distinct values. But there's no point in indexing a field that you never use in sorts, searches or queries.

To index a field, simply edit its Indexed property. In Access 2.0, you can also open an Indexes window (select View Indexes from within design mode) to view and edit all the table's indexes at once (Figure 3). This window also allows you to create multiple-field indexes.

Given that you can add and remove indexes at any time, it might make sense to remove indexes immediately before starting a large data-entry session, and to re-create them when you have finished.

9. Use Counter fields as primary keys

A primary key is particularly efficient if its data type is Counter. Use this data type for primary keys whenever you can.

10. Keep queries simple

You'll get the best performance from queries if you keep the selection criteria as simple as possible, and only include fields in the results that you really need. In particular, clear the Show checkbox for fields which are used in selection criteria but which do not need to appear in the results. If you want to sort the results, consider creating an index for the sorted field (see tip 8).

11. Optimise for Rushmore

Rushmore is probably the most worthwhile performance-related innovation in Access 2.0, and it's worth going to some trouble to use it properly. It is a proprietary technique which takes advantage of indexes to speed up queries. In some cases, it leads to dramatic performance gains, but you have to make sure that your queries are properly formulated to benefit from it. (FoxPro users will already be familiar with the concepts described here.)

In general, a query will be optimised for Rushmore if the selection criterion takes either of these forms:

<indexed field> <comparison operator> <expression>

or

<expression> <comparison operator> <indexed field>

For example, if Surname is an indexed field, then this expression is optimised:

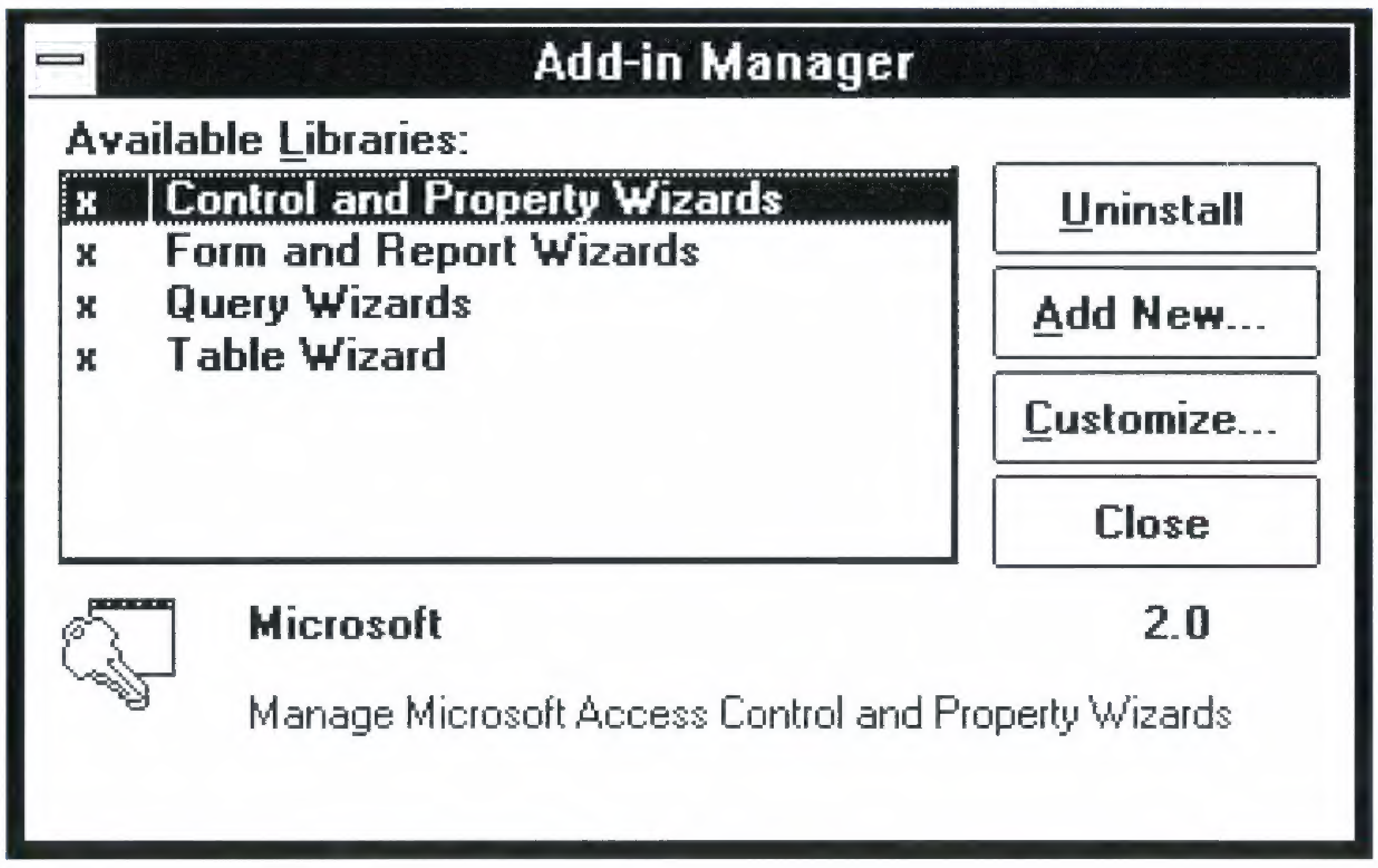
Surname > "K"

but this one is not:

Len (Surname) > 5

If the index is ascending, any valid comparison operator may be used. If it is descending, the operator must be "=".

If two or more optimised expressions are linked by an AND or OR operator, the resulting expression is also optimised. If the table has a multiple-field index, the query is optimised if the indexed fields are used in the order in which they appear in the Indexes window, starting with the first field and continuing with adjacent fields. For example, if the index is "Surname, Location", then an optimised query can be based on Surname, or on Surname and Location, but not on Location alone.



Manage Microsoft Access Control and Property Wizards

Figure 1 - Uninstall wizards to free up memory

> *"During the course of an Access session, the workspace quickly becomes littered with open tables, queries, forms and reports. Get into the habit of closing all open objects as soon as you can."*

The above rules apply both to the Criteria row in the Query window, and in the WHERE clause of a SQL SELECT statement. Rushmore works with native Access tables, and with attached FoxPro and dBASE files, but not with ODBC data sources.

12. Save query results

Suppose you frequently use a form or report based on the results of a complex multi-table query. By default, Access runs the query each time you open the form or report. This can take a long time, but it does ensure that you are always working with up-to-date data.

If the data is fairly static, or if you don't always require the latest version of it, it is more efficient to run the query once, save the results as a separate table, then use that table as the data source for the form or report. To do so, create the query in the usual way, then select Make Table from the Query menu. In the resulting dialogue, enter a name for the new table.

Next, go ahead and run the query. Having done so, invoke design mode for the form or report, and change its Record Source property to show the name of the new table. You can re-run the query from the main database window any time you want to refresh the data in the form or report.

13. Simplify forms and reports

You'll find that the processing of forms and reports is a lot faster if you keep their designs as simple as possible. Be stringent in reducing the number of fields, labels, frames, lines, rectangles and buttons on the form or report, and avoid using overlapping controls.

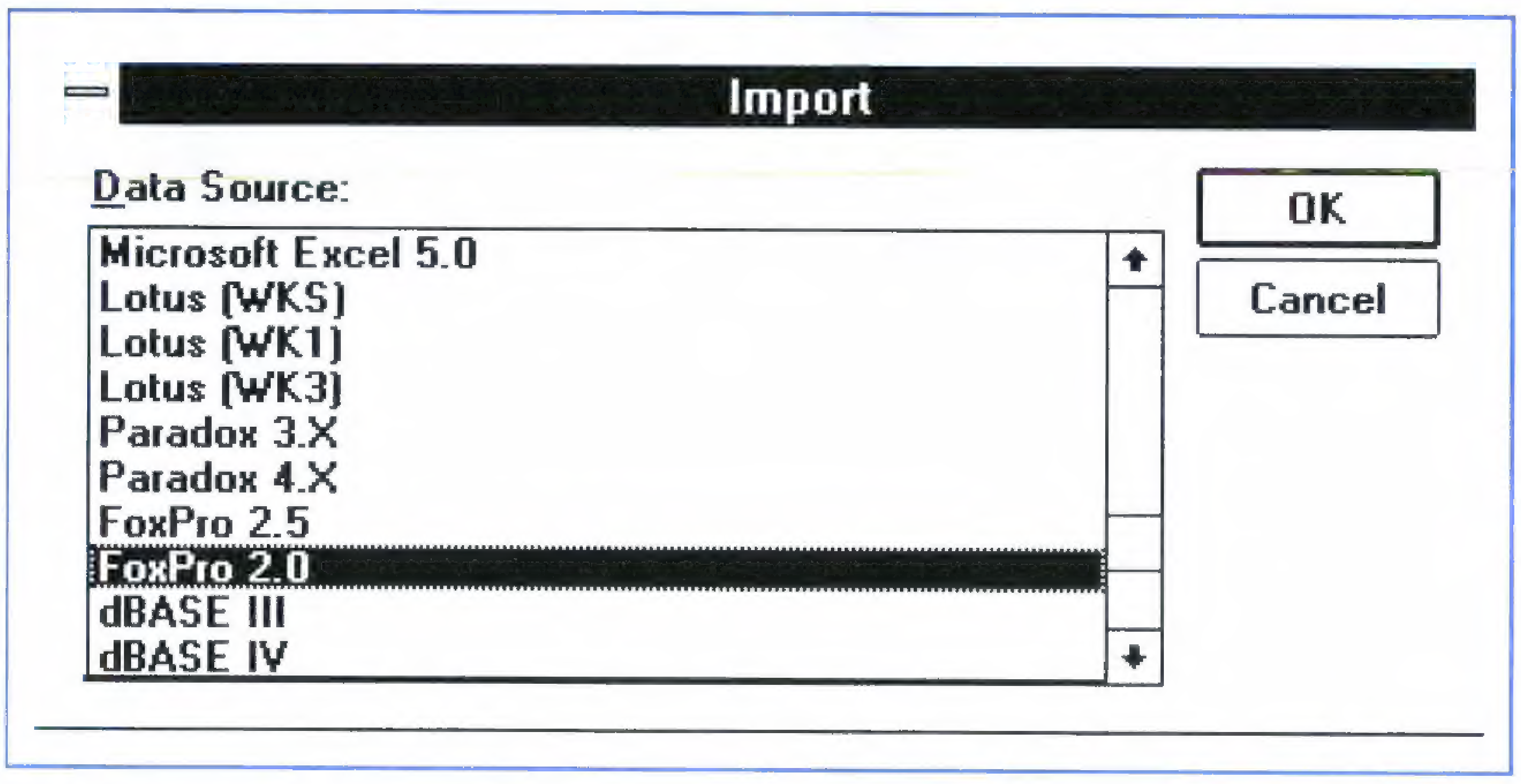
On command and toggle buttons, use text labels rather than bitmaps. If you must use bitmaps, keep them as small as possible, and use monochrome images rather than colour.

14. Use OLE sparingly

Access allows you to store OLE objects in tables, and to add OLE objects to forms and reports. This is a very handy feature, but it does have a detrimental effect on performance, so use it sparingly.

For Paintbrush, Graph and similar objects, you can improve performance by converting the object to a picture. To do so, select the object in design mode, then choose Change to Picture from the Edit menu. If the object is stored in a table, you will have to locate its record in form view in order to select the object. Once you have converted the object to a picture, any links with the original application will be maintained, but you will no longer be able to edit the object by double-clicking on it within Access.

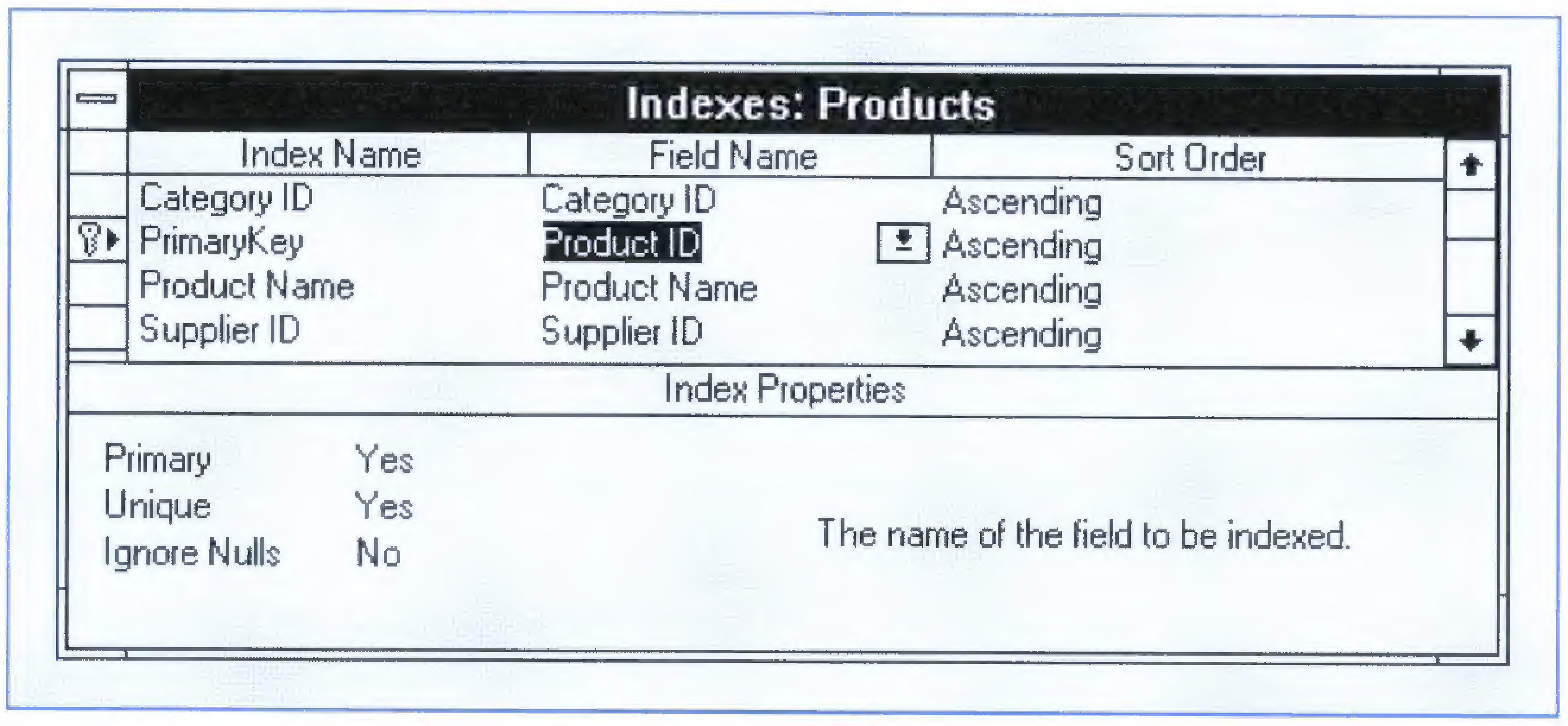
Another approach is to display the object as an icon. For example, if a form contains data from an Excel worksheet, it will be more efficient to display it as a worksheet icon rather than to show the data itself. If the user wishes to see the data, he or she can



Data Source:

- Microsoft Excel 5.0
- Lotus (WKS)
- Lotus (WK1)
- Lotus (WK3)
- Paradox 3.X
- Paradox 4.X
- FoxPro 2.5
- FoxPro 2.0
- dBASE III
- dBASE IV

Figure 2 - When working with external data, import rather than attach



Index Name	Field Name	Sort Order
Category ID	Category ID	Ascending
PrimaryKey	Product ID	Ascending
Product Name	Product Name	Ascending
Supplier ID	Supplier ID	Ascending

Primary Yes
Unique Yes
Ignore Nulls No

The name of the field to be indexed.

Figure 3 - Use the Indexes window to help manage the table's indexes.

> *"Wizards are a great boon, especially in designing forms and reports. But they consume resources at a fearful rate, even when not in use. If you can manage without these handy helpers, pull the plug on them."*

double-click on the icon to open the worksheet. To achieve this, select the Display as Icon option in the Insert Object dialogue.

15. Be careful with list and combo boxes

When placing a list box or combo box on a form, use a value list rather than a table or query as its data source. If you must use a query, make it an Access query (as opposed to a SQL SELECT statement), and keep the result table as small as possible. If you can, convert the query results to a table first (see tip 12).

16. Use fast laser printing

If you use a laser printer, you can speed up the printing of forms and reports by setting the FastLaserPrinting property to Yes. This tells the printer to use its built-in line-drawing function to print the rules, borders and rectangles on the form or report. Without this option, the rules, etc. would be output as bitmaps, which takes longer.

Virtually all modern laser printers have a line-drawing function. No harm will be done if you enable the property with a printer that does not have this feature, but there will be no benefit either.

17. Open databases exclusively

If you are working in a multi-user environment, open your database in exclusive mode whenever you can. This can lead to a significant performance gain, since it allows Access to dispense with its usual record locks. You will only be able to open a database exclusively if no other user is using it at the time. Once you have opened it in this way, no-one else will be able to open it (exclusively or otherwise) until you close it again.

If you only want to retrieve data without altering it (and without adding or altering forms, reports, queries etc), open the database in read-only mode. This, too, will speed up processing by eliminating some of the record locking.

Both options are available as checkboxes in the Open Database dialog.

18. Store data locally

If, on the other hand, you are working in a single-user environment, store Access itself, and your database, on a local drive if possible.

19. Check INI file settings

Experimenting with the settings in the [ISAM] section of the Access INI file can sometimes boost performance, though you might need to resort to trial-and-error to find the best values in each case.

The MaxBufferSize setting determines the amount of memory which Access reserves for its internal page cache. A high figure will speed up the processing of large tables and complex forms, but only by taking RAM from the overall memory pool. Experiment to find the setting that's best for you. The maximum figure allowed is 4096 KB, the default 512, and the minimum 18.

ReadAheadPages determines the size of the read-ahead cache, which is used for sequential reads of large tables. Bumping it up from its default setting of 16 (8 in Access 1.x) might be beneficial if you are processing large numbers of consecutive records, especially on a network. The maximum is 31 and the minimum 0 (the figure represents the number of 2 KB pages).

If you use Access 2.0, these settings are held in MSACC20.INI. Users of Access 1.x should look in MSACCESS.INI. Both files are installed in the Windows directory.

20. Upgrade to 2.0.

Finally, if you have not already done so, upgrade to the latest release of Access. Apart from offering a tempting range of new features, version 2.0 is generally more efficient than its predecessors. It comes with a new set of ODBC drivers which should speed up access to external data. And, of course, it brings you the considerable benefits of Rushmore (see tip 11).

In Conclusion

Not everyone will be able to take advantage of all the suggestions offered here. Some of the tips will only work in certain circumstances; others will involve trade-offs which you might find unacceptable. But most users should find several useful speed enhancers in the foregoing pages. Use them wisely and you might decide that Access is not such a bad performer after all.

PCSA

The Author

Mike Lewis is a Windows and database consultant and writer.

Understanding Optical Storage Technology

Dave Stott launches an investigation into the workings of optical storage products.

It appears that data, to coin a phrase, expands to fill the storage space available to hold it. Despite the fact that hard disk capacities have increased steadily over the last few years we seem to face a constant battle to find sufficient storage space for the data that we wish to keep on-line.

You would think that with 1 GB, 2 GB and even 4 GB SCSI hard disks available for PCs there would be more than enough storage capacity for most people. Indeed these high capacity hard disks are proving extremely useful. However, they still lack one important characteristic, they are not generally designed as removable media and, therefore, the user's total capacity is limited by the number and size of hard disk drives that can be accommodated by the system.

Databases with BLOBs (Binary Large OBjects), graphics files, digitised photographs, full motion video images, high quality sound and document scanning all consume vast quantities of that storage space. Combined with a general reluctance by users to throw any potentially useful data away means that the demand for more and more on-line storage space is constantly increasing.

So what is needed is a removable high capacity, relatively fast on-line storage medium which can be used to provide practically unlimited storage capacity. Fortunately the development of optical storage technology looks set to address these requirements.

In this article I will cover the various different types of optical storage available, explain the technology used and provide examples of the suitability of each type to meet specific needs. Although optical storage technology utilises different sized media (14 inch, 12 inch, 5.25 inch and 3.5 inch) I will only be dealing with the last two of these form factors as these are the main focus of current development. I shall also describe some of the latest developments which look set to increase the attractiveness of optical storage in the future. But let's start with a brief history of the technology itself.

Background

Optical storage is based on the use of the laser. When the laser was first invented in the early 1960s the scientific world had no idea that one of its main properties, the ability to be focused into an extremely small beam, would be an important factor in the future developments of computer technology.

By the early 1970s laboratory experiments were being conducted with optical recording materials which proved that the technology was capable of being used to record and retrieve data. Following on from this, a joint project by Philips and Sony led to the development of the Compact Disk in the early 1980s. This was originally used for the storage of digitised audio signals and was to become the precursor of the CD-ROM.

At around the same time, several companies were working on write-once recording techniques with the result that a number of different methods emerged. Ablative, dye-based, dual alloy, phase-change and thermal bubble recording techniques were all developed over a few years and they all utilise lasers to record information by irreversibly altering the reflectivity characteristics of a thermally sensitive material.

Since the mid 1980s research and development has concentrated on the area of rewritable optical storage technologies. The two major techniques employed to achieve this are magneto-optical recording and phase change recording, both of which are described in detail later.

CD-ROM

The advent of multimedia has triggered the adoption of the CD-ROM. CD-ROM is the most widely used optical storage technology around. As its name implies it is a read only storage media which has been developed from Compact Disc Audio technology.

A master disk is created with tiny pits recorded on it. From this master, CD-ROMs are pressed in the same way that audio CDs are manufactured. CD-ROM drives use a low powered laser to reflect off the pits on the disc's surface. These reflections are converted into electronic pulses and form a stream of data for subsequent processing.

Most CD-ROM discs are 4.75" format and the media is single sided offering approximately 540 MB of storage. More recently 3.5" format CD-ROMs have emerged which are designed to provide CD-ROM facilities on notebook PCs. These smaller discs currently offer a storage capacity of around 180 MB.

Early CD-ROM drives suffered from rather poor data access speeds. This was mainly due to the disk rotation speed. This was a major drawback when retrieving full motion video from discs. The amount of data requiring to be transferred to maintain a decent frame rate resulted a very jerky image. In order to address this problem manufacturers now produce double speed, triple speed and even quadruple speed drives. Most CD-

ROM drives utilise a SCSI (Small Computer System Interface) interface and the introduction of SCSI-2 has enabled much faster data transfer rates to be achieved.

The CD-ROM is primarily used for the distribution of large amounts of data and there are thousands of discs available ranging from shareware compilations to complete databases. It is also an ideal medium for software distribution and many software manufacturers offer their products on CD-ROM, indeed in the case of products such as OS/2 and Windows NT this is the preferred delivery mechanism.

Recordable CD-ROMs

Like audio CDs the scale of production has an important effect on the unit cost and often it can be uneconomical to produce small quantities of CD-ROMs. With this in mind, drives that enable the recording of CD-ROM discs have recently come on the market. These can be used to create small numbers of CD-ROMs on blank CDs. They are particularly useful for producing samples and enable testing of CD-ROMs prior to committing to large scale pressing.

Commonly known as CD-R this technology was rather expensive initially, with typical units costing in excess of $20,000. This high price was mainly due to the large hard disks incorporated in these early units. A large hard disk is required to enable the software to build an image of the target CD-ROM prior to transferring the data. Nowadays manufacturers have introduced CD-R drives without integral hard disk and this has dramatically reduced prices to the region of $4,000. However, these new drives which use SCSI interfaces, assume that you have suitable high capacity hard disk on which to generate the CD-ROM image.

One such drive is the new CDD521 from Philips. This can be used to record CD-ROM(XA), CD Multi-Session, CD-I and CD Audio discs. Windows-based software is also available to make the production of small quantities of CD-ROMs very straightforward. The system can produce CD-ROM disks which conform to the ISO9660 format standard which means that the resulting disks can be read by both IBM PC and Apple Mac computers.

WORM

Unlike CD-ROMs on which users cannot record data the WORM (Write Once Read Many) is specifically designed as a non-erasable recording media. This makes it ideal for tasks such as archiving and maintaining audit trails where data security is paramount.

The WORM media consists of an active layer embedded between two transparent substrates typically made from polycarbonate or glass. The WORM drive uses a laser operating at two different power levels, high power for writing data and low power for reading data. The recording of data can be achieved by a variety of techniques in which the laser is used to produce tiny pits, spots or bubbles on the active layer. These small changes to specific areas of the active layer alter the reflectivity when compared to the normal background state of the disk. This difference in reflectivity is then read by the low power laser in order to retrieve the data.

One very important aspect of the WORM disk is its ability to reliably retain information over long periods of time especially if they are to be used for data archiving. However, as with all new data recording methods, it is impossible to prove the long term reliability of such media. Therefore, accelerated aging techniques have been used by manufacturers to provide an estimated data storage life.

At present there are five methods of

	Media Capacity	**Access Time**	**Removable Media**	**Rewritable Media**	**Autochanger Available**
Fixed Magnetic Disks	20 MB - 4 GB +	8-65ms	NO	YES	NO
Removable Hard Disks	60 MB - 200MB	15-30ms	YES	YES	NO
Bernoulli Box Systems	20-90MB	65ms	YES	YES	NO
Floppy Disk Systems	360KB - 21MB	160-350ms	YES	YES	NO
1/4" Magnetic Tape	20MB - 1.3GB	minutes	YES	YES	NO
8mm Magnetic Tape	1GB - 8GB	minutes	YES	YES	YES
4mm Digital Audio Tape	1.3GB - 2GB	minutes	YES	YES	YES
VHS type Data Recorders	2.5GB - 10GB	minutes	YES	YES	NO
5.25" WORM Disks	230MB- 1.28GB	50-300ms	YES	NO	YES
5.25" M-O Disks	256MB- 1.3GB	25-110ms	YES	YES	YES
3.5" M-O Disks	128MB-230MB	25-50ms	YES	YES	YES

Figure 1 - Comparison of Selected Magnetic and Optical Storage Technologies

Optical Storage

generating the recorded data on different types of WORM discs:

Tellurium Thin Films: This is the most common approach to WORM recording. The disk consists of a thin film of tellurium and other stabilising materials such as selenium or lead covered with a transparent protective layer. The laser creates microscopic holes or pits in the thermally sensitive layer. This process is termed ablation (from the Latin ablatio - to carry away) recording. Estimated data storage life 30 to 40 years.

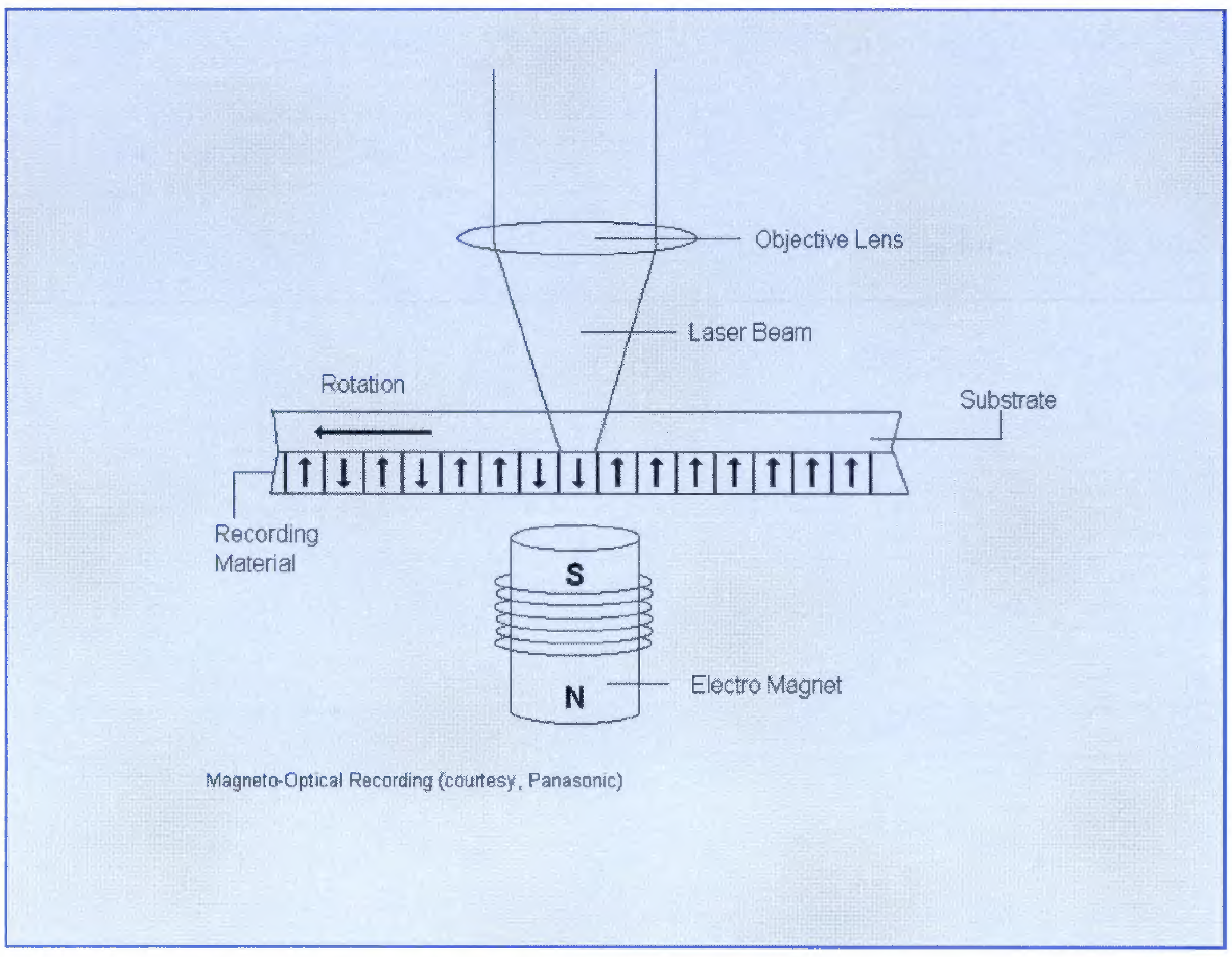
Dye Polymer Media: This approach uses a transparent polymer layer with an infra-red absorbing dye coated on a plastic or glass substrate. Recording is achieved by matching the laser with the dye's absorption wavelength. This generates heat that forms pits or bumps of different reflectivity on the surface. Estimated data storage life 15 years.

Dual Alloy Media: A layer of tellurium-bismuth alloy is sandwiched between two layers of selenium-antimony alloy. The laser fuses these together to form a combined four element alloy which has a different reflectivity from the surrounding area. At present this recording method is employed exclusively by Sony. Estimated data storage life 100 years.

Phase-Change Recording: Here tellurium or selenium based alloys are again heated to a certain temperature by the laser as in the ablation method described above. However, rather than actually removing material, phase-change recording relies on altering the state of the layer between crystalline and amorphous. Each of these states demonstrates different reflectivity. Panasonic use this recording method for its WORM devices. Estimated data storage life 15 to 50 years.

Thermal Bubble Recording: Laser heat is used to evaporate a polymer layer to selectively form bubbles in a thin layer of gold or platinum. The bubbles open to form pits that reveal a reflective underlying layer. Although this was one of the first techniques used for recording WORM data it has not proved very popular and is now only used by two or three manufacturers. Estimated data storage life 30 years.

Each of the above methods basically achieves the same thing. However, it is important to remember that the recording method used is drive and media dependent and, therefore, discs recorded by a particular method cannot subsequently be used on a drive using a different technology.



Magneto-Optical Recording (courtesy, Panasonic)

Figure 2

The most popular WORM format uses a 5.25 inch disk size with a typical average access time of around 50 milliseconds. Currently capacities are in the region of 650 MB per disk although some manufacturers have pushed this to 1.28 GB per 5.25 inch disk.

It is worth noting that with the exception of CD-ROM all optical storage discs are generally double sided and the vast majority of drives cannot access both sides without having to turn the disk cartridge over. However, it is quite common for a manufacturer to state that a particular drive has, say, a 512 MB capacity. This in fact means that the disk has two sides with 256 MB capacity on each side. Therefore, only half of the drive's stated capacity is available on-line at any one time.

Read/Write Optical disks

Although the WORM device has its own particular advantages for archiving data, it is less suitable for more general mass storage requirements. Most data is not permanent and users often need the ability to make changes to it. This fact has led manufacturers to develop the read/write optical disk.

These offer much more flexibility and are an ideal media for supplementary on-line storage. There are two main recording technologies used for read/write optical disks; Magneto-Optical (M-O) and Phase Change.

Magneto-Optical

Magneto-Optical is the more popular recording method used for read/write devices. All M-O devices utilise two different aspects of Physics in order to work. As with magnetic media in general, a layer of metal or oxide provides the recording surface. At the microscopic level small areas of this material (normally referred to as magnetic domains) are used to store the data in binary form.

Whilst investigating the effect of heat on various magnetic materials it was discovered that at a certain temperature (known as the Curie Point) a heated magnetic domain will align its polarity to match the polarity of an adjacent magnetic domain. Using these physical phenomena it is possible to arrange the domains on a magnetic surface so that their poles are aligned (eg, north pole up or south pole up which equates to 0 or 1 - see Figure 2). The actual magnetic domains are only about one micron in diameter (ie, one millionth of a metre) and consequently extremely high recording densities can be achieved.

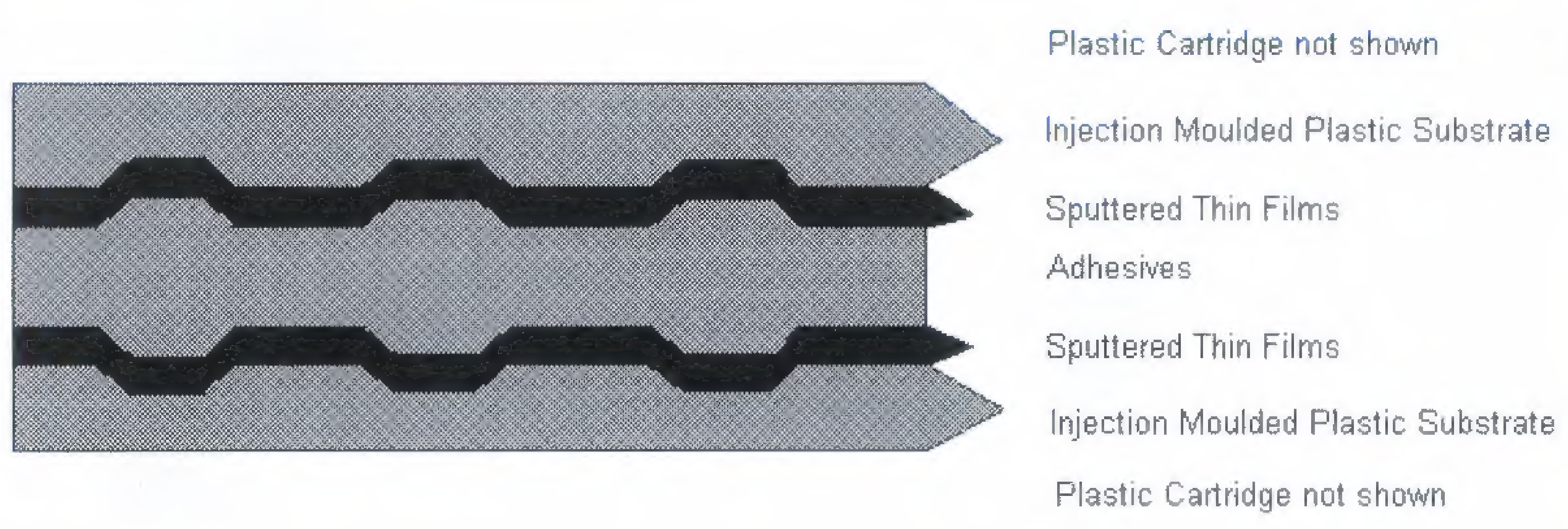
The heat for the process is provided by a primary laser which is focused by lenses to act directly on the magnetic surface. This laser raises the temperature of small areas of the disk to approximately 180 degrees C. The magnetic field (known as a bias field) used to change the polarity of the magnetic domain is produced using a tiny electro magnet underneath the disk which is electrically switched to alternate its polarity.

In order to work, M-O drives use a two phase writing technique to store and erase data. The first phase will align the polarity of a series of magnetic domains in one direction only, thereby setting the individual bits within a sector to 0. The next phase will only change the polarity of certain domains, changing some 0s to 1s thereby recording the relevant information. Obviously this two phase approach requires two complete revolutions of the disk and consequently has an effect on the overall performance of the drive.

A secondary, less powerful laser beam is used for reading the data previously written on the disk. This relies on another physical characteristic of Magneto-Optics known as the Kerr Effect. Here the interaction of the laser light with the polarity of a particular magnetic domain is crucial. The light reflected back from a certain magnetic domain is modified by the actual magnetism and the reflected laser light is rotated very slightly either clockwise or anti-clockwise depending on whether the source of the reflection is a north pole or a south pole. The actual amount that the reflected laser light is rotated is only about 1 angular degree and a sensitive detector is used to determine the direction of rotation. In this way the data is read from the disk. It must be borne in mind that both of these processes (reading and writing) are carried out whilst the disk is rotating at up to 3,600 rpm. This fact coupled with the microscopic scale gives an indication of the level of accuracy involved and the consequential need for precision engineering.

The disks themselves consist of two sides composed of multiple layers of different materials (see Figure 3). The magnetic layer most often consists of alloys of terbium, iron, cobalt and traces of other elements. Two of these layers are bonded together onto an adhesive layer and the complete sandwich is covered with either a moulded plastic substrate or encapsulated in a special type of glass

M-O disks and drives are available in both 5.25 inch and 3.5 inch form factors. The 5.25 inch disk provides up to a maximum of 1.3 GB storage with an average access time of around 50 milliseconds. Most 3.5 inch disks offer a 128 MB capacity although there are a few 3.5 inch drives around which can handle higher capacity disks. One such drive is the new Fujitsu M2512A which can support 230 MB capacity 3.5 inch disks and has an average access time of less than 35 milliseconds.


Plastic Cartridge not shown
Injection Moulded Plastic Substrate
Sputtered Thin Films
Adhesives
Sputtered Thin Films
Injection Moulded Plastic Substrate
Plastic Cartridge not shown

Magneto-Optical Media (courtesy, Hewlett Packard)

Figure 3

Phase-Change

The Phase-Change (P-C) technology used for read/write media is very similar to the P-C WORM described above. The major difference between the two is in the material which is used as the active recording layer. This material has the property of reversible phase change, and unlike WORM disks it starts off with a completely crystalline (high reflectivity) structure.

A very high powered laser is used to locally heat the active layer wherever a data bit is to be recorded. This melts the crystalline structure which then rapidly cools to form an amorphous (lower reflectivity) spot. By reheating an amorphous spot with a slightly lower powered laser (to a temperature below the melting point of the active layer), recrystallisation will occur, and the spot will revert back to its original crystalline structure (high reflectivity). This means that laser beam power alone (and subsequent temperature generated) can be used to change the active layer between the two states, crystalline and amorphous.

In addition this approach means that unlike M-O technology the P-C method can write data in a single pass (see Figure 4). It uses the combination of the two laser power levels to effectively switch spots to their opposite state. Therefore, P-C drives are generally twice as fast as M-O drives during data writing operations.

At present only Panasonic use P-C technology to any significant degree and all their drives adopt the 5.25 inch form factor. The disks available have a maximum capacity of 1 GB and the fastest Panasonic drive has an average access time of 90 milliseconds.

Multifunction Drives

A special group of rewritable optical drives, termed multifunctional drives, can operate in either rewritable or write once mode. These were first

Optical Storage

introduced in 1990 and are proving to be very popular as a single drive can be used for two different purposes.

The majority of multifunction drives use different media for each operation. The drive is capable of detecting the type of media inserted and automatically configures itself for the appropriate recording method. In order to protect the original data on a write once disk the drive's command set will be altered so that it does not support data destructive operations such as file deletions or media reformatting.

Some multifunction drives utilise different recording methods for dealing with rewritable and write once media. For example, the LaserDrive 520 from Laser Magnetic Storage International (4425 Arrows West Drive, Colorado Springs, CO, USA - Tel: 719-593-7900) uses ablative recording of tellurium based write once media and magneto-optical recording technology for rewritable media. Similarly the Pioneer DE-7001 drive uses magneto-optical recording in its rewritable mode, but dye based media for write once recording.

Another approach is to use the same recording technology for both purposes. Panasonic's LF-7010E drive uses phase change recording but uses slightly different disk media for each type of operation. This begs the question, how do manufacturers ensure that a write once disk cannot be overwritten?

Hewlett Packard offers a multifunction optical disk drive which uses magneto-optical media for both write once and rewritable recording, although different disks are required for each operating mode. Write once and rewritable media are identified by a factory inscribed code which are detected by the multifunction drive when a given optical media disk type is loaded. The drive automatically invokes the recording mode appropriate to the loaded disk.

As noted above, the multifunction drive's write once command set does not support deletion or overwriting of data. Accidental erasure is, therefore, impossible. Additional coding prevents recording in data blocks which have already been used. In addition, write once and rewritable media are colour coded and labelled differently in order to distinguish between the two types.

Flopticals

Despite its name the Floptical is not strictly an optical storage medium. It uses the same magnetic recording and reading method as a standard floppy drive. Developed by Insite Peripherals the Floptical drive uses a low powered laser to maintain extremely accurate positioning of the read/write heads. In this way much higher recording densities can be achieved than with conventional floppy disks.

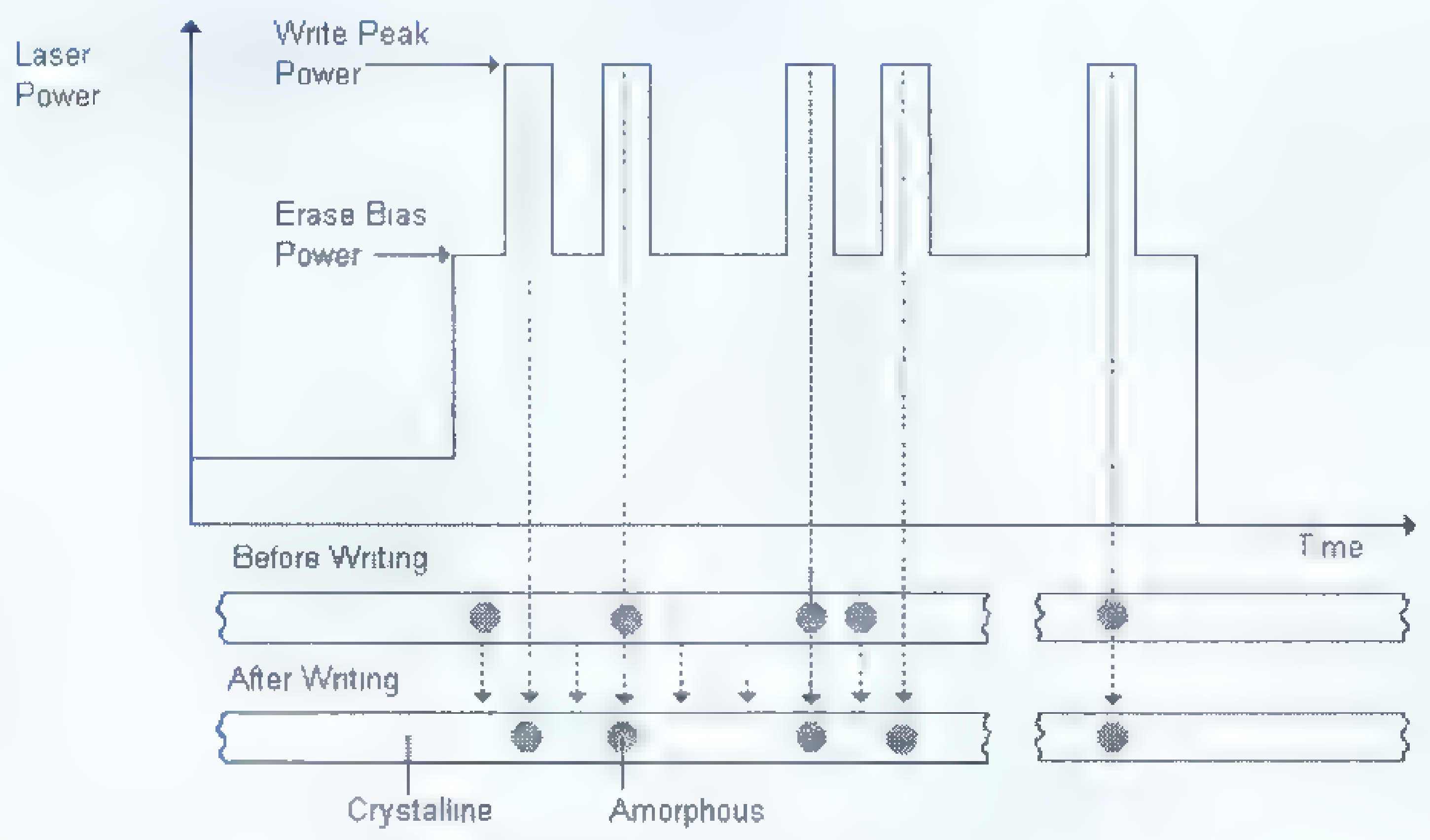
The Floptical drive uses basically the same type of 3.5 inch diskettes as a standard floppy drive except that they feature Barium Ferrite as their recording material. A typical Floptical diskettes can hold a maximum of 21 MB of data and they are mainly used as a form of data backup where they can offer a relatively quick and cheap solution.

Floptical drives generally use a SCSI interface and transfer rates are up to three times faster than a normal floppy disk drive with average access times around 65 milliseconds. Another significant advantage is that Floptical drives are also backwards compatible and can be used to read and write all existing 3.5 inch HD 1.44 MB and DD 720 KB standard floppy disks.

Autochangers & Jukeboxes

Despite the fact that optical disk drives offer relatively large storage capacities the fact that both sides of the disk cannot be accessed without turning the disk media over is seen as a significant drawback. At best this is a little inconvenient and at worst if there is data stored across a large number of disks then it can be almost totally impractical. A few manufacturers have produced drives which can automatically access both sides of the disk by using dual read/write heads as with floppy disk drives. However, the vast majority of drives available still use a single read/write head.

One solution to this problem is to utilise multiple drives to make a greater quantity of data available online at any one time. Several optical storage equipment manufacturers offer preconfigured models with twin drives housed in a single chassis. The Ricoh RS-8200HH, for example, is a



Phase Change - Direct OverWrite Recording (courtsey, Panasonic)

Figure 4

dual drive optical storage subsystem that provides on-line access to 800 MB of data recorded on WORM disks. The dual drives are available in either horizontal or vertical casings and use Ricoh's dye based, 5.25 inch WORM media which can store 400 MB per recording surface.

Alternatively you could set up a system using multiple individual optical drives. The fact that most manufacturers have opted for SCSI interfacing for their equipment makes this option relatively easy to implement. However, the cost of such a multi drive solution may prove to be prohibitively expensive.

A more practical and cost effective solution for the provision of very large amounts of on-line storage is the autochanger or "jukebox" device. In essence, an optical disk autochanger is a mass storage peripheral that can provide access to gigabytes or even terabytes of data held on multiple optical disk cartridges without the need for operator intervention. Normally an autochanger is a floor standing unit although several manufacturers are now producing desktop versions.

Regardless of size, an autochanger consists of at least one optical drive and a robotic mechanism for switching disks held in stacks or storage bins. Generally autochangers are subsystems which have their own dedicated CPU and memory that receives and processes commands from the host system. Once an instruction is received to access data held on a specific disk, the robotics will remove any disk that is currently in the drive(s) and insert the correct disk for use. Obviously in order to work efficiently the host system must be made aware of what discs are available. This is usually achieved by use of corresponding mapping tables held both within the host and the autochanger itself. If any of the disks held within the autochanger is removed, replaced or added to then the mapping information has to be updated prior to using the subsystem. Some autochangers also include an import/export facility for the manual addition and removal of disks.

Whilst they operate as on-line peripheral devices, optical disk autochangers are best described as providing "nearline" access to stored data, since requests for certain data may result in a disk having to be swapped into a drive and causing a slight delay in access. Access times for autochangers are consequently measured in seconds rather than the milliseconds normally associated with true on-line storage devices. Typically interchange times for optical disks range from six to 20, with newer models usually providing faster access than older versions. Realistically "nearline" storage should be viewed as secondary storage and used as a way supplementing primary storage media such as fast hard disks.

To improve disk interchange times some autochangers are configured with two optical disk drives as standard equipment. This allows a previously mounted disk to be removed from one drive at the same time that a specified disk is inserted into the other drive. In other cases a given autochanger can be configured with multiple drives. Depending on the model, between two and five drives may be supported. However, because the drives occupy space that could otherwise be used for additional disks, additional drives invariably reduce the overall storage capacity of the device. This means that there is frequently a trade-off between speed of access and capacity.

> *"There is a strong possibility that in the future WORM optical technology will replace magnetic tape as the preferred medium for archival data storage."*

Depending on the model, 5.25 inch autochangers may be configured with write once or rewritable drives and media. Several autochangers are equipped with multifunctional drives or with a combination of write once and rewritable drives. Described as multifunctional jukeboxes, such peripherals can store both WORM and rewritable disks. When a given disk is requested, the robotic mechanism will automatically load it into the appropriate drive.

Overall capacities for autochangers vary enormously. At the low end there are the small desktop models which generally have a single drive and can accommodate up to five 5.25 inch disks. A typical desktop model would offer a maximum of 11 GB of total storage and would cost in the region of $6,000 to $9,000 including media and software. Top of the 5.25 inch form factor range is the recently launched Hewlett Packard 200T Optical Disk Library. This comes with two multifunction drives as standard and can be upgraded with a further two drives if necessary. It can hold a maximum of 144 1.3 GB discs giving a massive overall capacity of 188 GB. A fully configured HP 200T would set you back around $90,000 including the media and software.

Larger capacity jukeboxes are available that are based on either 12 inch or 14 inch media form factors. These can generally provide several hundred gigabytes of storage and some can even reach low end terabyte figures.

The problem of providing access to multiple CD-ROM discs has been addressed by several manufacturers. The most popular device is the Pioneer DRM-6000 CD-ROM minichanger which is based on audio CD autochanger technology. This uses interchangeable cartridges which can be preloaded with up to six discs. A single cartridge can provide access to approximately 3.2 GB of data. Up to seven minichangers can be attached to a single SCSI controller for a total capacity of 22.4 GB per configuration. Such devices are an ideal way of providing access to multiple CD-ROMs especially on a network. They have also been successfully utilised by on-line Bulletin Board Services (BBS) to

Optical Storage

provide, for instance, access to vast amounts of shareware.

Other Optical Devices

Despite the fact that direct access disk is the is the predominant optical storage technology in use today there are two other types of recording media; optical cards and optical tape. Of these two, optical cards have been established the longest.

Based on technology developed by Drexler Technology Corporation the LaserCard is the best known and most widely demonstrated optical card. The LaserCard measures approximately 3.4 x 2.1 x 0.03 inches, roughly the size of a credit card. It is based on a proprietary, patented optical recording material called Drexon. Drexon features metallic silver filaments dispersed in a gelatin matrix which is coated on glass, plastic or other substrates. The LaserCard's capacity is around 2.86 MB so it could not be termed a mass storage media. Even after extensive trials over a number of years the optical card has not proved very popular. Nevertheless several Japanese companies are still experimenting with optical card technology.

A more viable proposition is optical tape. This has been developed over a number of years as a method of providing huge mass storage capacities. As its name suggests, optical tape is a flexible, ribbon shaped substrate coated with an optical recording material. The optical recording material is Digital Paper, which was developed by Imperial Chemical Industries (UK), and consists of infrared sensitive, dye polymer coated on a flexible polyester substrate.

At present CREO Systems (Canada) are the only company actively supplying optical tape drives, marketed by Honeywell, for commercial use. The CREO 1003 tape drive employs a 35 millimetre wide tape on a 12 inch reel with a length of 880 metres. The drive uses a multiple laser diode array to read and write multiple tracks of data in 32-bit increments. Its recording capacity is 1 terabyte of data per reel, the equivalent of about 5,000 reels of standard 9 track magnetic tape recorded at 6,250 bits per inch or nearly 100 of the highest capacity WORM disks.

Optical tape is still in the early stages of development and, as with the first generation of optical disks, optical tape technology is at present confined to write once capabilities. However, there are a number of companies exploring the feasibility of magneto-optical tape including 3M and Schlumberger, and although still at the experimental stage the potential for M-O tape storage is enormous.

Application Areas

There are two main factors which make optical storage attractive compared to conventional magnetic media, the capacity for holding large amounts of data and the long term "shelf life" and reliability of recordings. Any computer processible information that can be recorded on magnetic media can be stored on optical discs as an alternative. But WORM and rewritable optical technology have specific strengths and weaknesses which make each of them particularly suitable to certain applications.

By their very nature WORM systems are especially suited for use in areas which require the long term archiving of data. Data such as audit trails and historical records are ideal candidates for WORM media. This data is unlikely to change much, if at all, over a long period of time. Traditionally magnetic tape has been used for this purpose and previously it has proved to be a very cost effective archiving media. However, magnetic tape can deteriorate dramatically over a period of just five or six years and, therefore, despite the additional media cost, WORM technology is proving to be more appealing. Similarly, the use of Computer Output to Microfilm (COM) for archival purposes is being challenged by Computer Output to Laser Disk (COLD) systems.

Graphics, document images, video images and to some extent high quality digital sound recordings, all consume vast amounts of storage space. To date the single most popular application to embrace optical storage significantly has been Document Image Processing (DIP).

Computer configurations that generate, store, display, retrieve, print, disseminate and process digitised document images are the basis for DIP. Such systems may utilise magnetic or optical media for document storage. The large number of bits associated with digitised document images necessitates a high capacity storage medium. As an example a 3.5 inch optical disk with 128 MB capacity can, with suitable compression techniques, store approximately 1,000 pages at a resolution of 400 dpi. A 1.2 GB 5.25 inch disk could store up to 9,600 pages at the same resolution.

The rapid proliferation of LANs has stimulated considerable interest in optical disk drives and autochangers as high capacity data storage devices. In theory, any optical disk drive can be attached to a network file server, provided that the drive is compatible with the server's hardware configuration and operating system. Recognising the enormous potential for network applications, an increasing number of

> *"In order to work, M-O drives use a two phase writing technique to store and erase data. The first phase will align the polarity of a series of magnetic domains in one direction only, thereby setting the individual bits within a sector to 0."*

> *"At present only Panasonic use P-C technology to any significant degree and all their drives adopt the 5.25 inch form factor. The disks available have a maximum capacity of 1 GB and the fastest Panasonic drive has an average access time of 90 milliseconds."*

systems integrators and VARs are certifying their optical storage subsystems for use with specific combinations of LAN hardware and software.

On The Horizon

After their commercial introduction the capacity of 5.25 inch optical storage devices remained static at around 650 MB for about four years. It was only during 1993 that 1.3 GB M-O drives became available. However, due to certain technological advances there is likely to be a dramatic increase in capacity and performance over the coming years, eventually leading to 10 GB plus capacities by the latter part of 1998.

The first big leap is likely to come from the adoption of Pulse Width Modulation (PWM) as opposed to the Pulse Position Modulation (PPM) currently used. PPM relies on individual dots to determine 0s or 1s whilst PWM uses the transition from one state to another at the leading and trailing edges of a dot to signal a change from 0 to 1 or 1 to 0. This means that sequences of consecutive 1s or 0s can be determined by time and thus be recorded in a smaller area. The use of PWM coupled with the reduction in track width from the current 1.39 microns to 1.15 microns will enable 2.6 GB capacities to be achieved.

Researchers at Sony are working on a technique with lasers that are capable of Magnetic Super Resolution (MSR). Rather than relying on the focusing capabilities of the laser itself MSR uses a very small aperture to reduce the effective track width to around 0.3 microns. This approach should result in 5.25 inch disks capable of storing a maximum of 5.2 GB of data. However, at such small track widths the problem of cross-talk interference often arises and this is being addressed by using cross-talk cancellation techniques. Here three tracks are read simultaneously and a custom chip acts as a filter to eliminate the unwanted signals.

By using additional layers in the media, researchers are beginning to tackle the two pass writing drawback inherent in the present design of M-O devices. Direct OverWrite (DOW) will put M-O drives on a par with Phase-Change technology and should eventually mean faster data access to the point were M-O devices begin to challenge conventional hard disks.

To achieve capacities of more than 10 GB researchers are investigating the use of lasers with shorter wavelengths. Both blue and green laser diodes operate at wavelengths much shorter than the normal red lasers in use today. However, in order to utilise such blue or green lasers new recording materials will have to be developed which can respond to the associated wavelengths generated.

Standards

Despite the fact that both the International Standardisation Organisation (ISO) and the American National Standards Institute (ANSI) are actively involved in defining and promoting standards for optical storage there are, as yet, no definitive standards for the technologies as a whole. Manufacturers have tended to adopt their own set of standards which means that, whilst disks and drives will generally be compatible within a particular manufacturer's product range, the same is not always true when dealing with equipment and media from different suppliers.

Sadly this means that the interchangeability of optical disks is not always possible and, therefore, you should be very careful in establishing such compatibility if the question of interchangeability arises.

Conclusion

After a somewhat slow start, optical technology appears to have finally established itself as a viable approach to the problems of mass data storage. As the demands for more and more data to be kept on-line inevitably increase manufacturers will continue to refine the technology to meet users expectations. Both media and more importantly equipment costs have fallen steadily making optical storage in general an attractive proposition.

There is a strong possibility that in the future WORM optical technology will replace magnetic tape as the preferred medium for archival data storage. Rewritable technology is already challenging conventional magnetic disks in terms of capacity but whether it can eventually start to match the access speed of modern hard disks remains to be seen.

PCSA

The Author

Dave Stott is a freelance writer and PC specialist with an IT industry background.

Understanding Colour Notebook Technology

Mike Lewis explains the different screen technologies used in colour notebook PCs.

Once users become accustomed to high-quality colour on their desktop PCs, they hate having to settle for less on their portables. But choosing a colour notebook involves some difficult trade-offs. Do you opt for the low cost and easy availability of a passive-matrix model and put up with a dull, washed-out image which has be viewed at exactly the right angle? Or do you choose the razor-sharp clarity of an active-matrix display, with its high price and long delivery time?

This article explains the various technologies used in colour notebook displays, and highlights the strengths and weaknesses of each. It won't necessarily assist you in choosing the ideal machine, but it will help you to evaluate the options, balance the trade-offs and give sensible advice to your users.

The Basics

All notebook computers use the same liquid crystal display (LCD) technology that has been around for over 20 years in watches and calculators. To understand how it works, think back to the school science experiment where you take two polarising lenses - such as those found in Polaroid sun-glasses - and hold them in such a way that you can look through both at the same time. Normally, the surrounding scene is visible through the two layers. Now twist one of the lenses through 90 degrees. Because the lenses only permit light to pass through in one plane, the image disappears.

An LCD consists of two of these polarising filters with a special crystal substance sandwiched between them. One of the filters is rotated through 90 degrees, which would normally prevent any light passing through. But the crystal has the ability to twist the plane of polarisation through 90 degrees, and therefore cancels out the effect of the filters. However, the crystal loses this ability when an electrical charge is applied to it. The result is a gate. When a charge is present, the gate is closed and no light passes through; when the charge is removed, the gate opens and light is again permitted to pass.

You can think of an LCD screen as an array of thousands of these gates. Each gate, or "shutter", switches one pixel on or off, thus forming the image. To incorporate colour, you simply add red, green and blue filters to each shutter, just as in a desktop CRT monitor.

Early LCD screens relied on ambient light to illuminate the image. Light passed through the display from outside and was reflected back out again. Though inexpensive and easy to make, these screens were inefficient, and only worked properly in ideal lighting conditions. Modern LCD displays have their own light source, either at the side or the back of the screen. In general, backlit systems produce better images than sidelit ones, but tend to be bulkier and require slightly more power.

Passive Matrix

The most important factor affecting the quality of an LCD image is the method used to toggle the gates. The simplest system is called passive-matrix. These screens incorporate two layers of parallel transparent strip electrodes. In one layer, the electrodes correspond to the columns of pixels on the screen. The other layer, which is placed on the other side of the liquid crystal element, has an electrode for each row of pixels, running at right angles to the column electrodes. To switch on a given pixel, the controller chip applies a charge to the column and row electrodes that intersect at that point.

This charging takes place on a cyclical basis. The controller starts at the top row of the display, charging each row in turn, along with the corresponding columns. It typically takes one-sixtieth of a second to cycle through all the rows. A standard VGA screen has 480 rows, which means that a given row only receives a charge for 1/480th of the cycle. This results in a dim, low-contrast image which is hard to see unless viewed at just the right angle.

> *"The secret of the success of active-matrix is that each pixel is controlled by its own transistor. Manufacturers use thin-film transistor (TFT) technology to place an array of tiny transistors right on the surface of the inner screen."*

To improve the image, manufacturers have resorted to using thicker crystal layers which twist light by 270 degrees rather than 90 degrees. This "supertwist" material enables the pixel to stay illuminated longer, so giving a slightly brighter display and sharper contrast. Even so, images still look dull and washed out, with contrast ratios of around 15 to 1 being the best that can be achieved.

Another problem with passive-matrix displays is the shimmering effect caused by the slow refresh cycle. To avoid this, manufacturers have to use a crystal substance with a low response time - about 300ms. That's much too slow for applications that involve any type of movement. Even keeping up with a moving mouse pointer in a Windows application can prove difficult for these displays.

> *"Virtually all notebook vendors have passive-matrix colour models in their catalogues, and these typically sell for only $300 to $600 more than their monochrome equivalents."*

Glossary

Active addressing
A system, currently under development, in which each row of pixels is controlled by an application-specific integrated circuit, which distributes a pattern of voltages across the intersecting columns.

Active-matrix
A system in which each pixel is controlled by its own transistor, these being mounted on a transparent film.

Backlighting
A method of illuminating the display, using tiny bulbs at the back of the screen.

Birefringence
A green or blue tint, caused by the light passing through the crystal being split into two components, polarised at right angles to each other. To overcome it, an extra layer of crystal is used to twist the light in the opposite direction.

Dual-scan
A variation of passive-matrix, in which the 480 rows (of a VGA display) are divided into two 240-row sections, each of which is controlled independently.

Passive-matrix
A system in which pixels are controlled by means of a grid of intersecting strip electrodes.

Response time
The time taken for a pixel to switch between states. A low response time causes a moving image to blur or disappear.

Sidelighting
A method of illuminating the display, using tiny bulbs at the side of the screen.

Supertwist
The use of a thicker crystal, which twists light by 270 degrees rather than 90 degrees. This allows the pixel to remain charged for longer, so increasing brightness and contrast.

Against these drawbacks, passive-matrix displays have the overriding advantages of low cost and easy availability. The technology is relatively simple, the screens are easy to make, and reliability is high.

Virtually all notebook vendors have passive-matrix colour models in their catalogues, and these typically sell for only $300 to $600 more than their monochrome equivalents (all prices in this article are in US dollars, and are based on street prices prevailing in the USA in the first half of 1994). Most users would probably find passive-matrix acceptable for straight text work, but would hate you for recommending one for graphics applications or Microsoft Windows.

Active Matrix

At the other extreme, active-matrix displays deliver images of exceptionally high quality. They are sharp, clear and uniformly bright, even when viewed from an angle.

They have no difficulty keeping track of moving objects such as the mouse pointer, and the better ones even support full-motion video. The downside is high prices and long delivery times.

The secret of the success of active-matrix is that each pixel is controlled by its own transistor. Manufacturers use thin-film transistor (TFT) technology to place an array of tiny transistors right on the surface of the inner screen. Each transistor acts as a gate for controlling the charging of the pixel. The charge stays in force throughout the refresh cycle, so the display suffers none of the quality defects of the passive-matrix system.

Colour Notebook

"If you want the very highest display quality available, with sharp, bright images easily viewable from any angle, capable of supporting animation and full-motion video, don't settle for anything less than active-matrix."

Because all the pixels can be controlled simultaneously, the screen lets more light through than does a passive-matrix model. This leads to a brighter display, with improved contrast and superior colour saturation. Contrast ratios as high as 60 to 1 are typical. Also, because the image is free of oscillation, a faster crystal material can be used, giving response times of 20ms or better.

Against these advantages, active-matrix displays tend to be bulkier and heavier than passive-matrix screens, usually adding around 500 gms (1.1lbs) to the overall weight of the notebook. They also consume more power, though this is partly offset by the fact that you don't need to keep the brightness turned up so high. For most users, these snags are easily lived with. People who have tried active-matrix computers never want to go back to the feeble look of a passive-matrix model.

The real problem with the technology lies in its sheer complexity. A 640 x 480 screen with three colour filters requires an array of 921,600 transistors. A very sophisticated manufacturing process is needed to produce this, and the reject rate is high - around 90%, according to some estimates. Not surprisingly, prices have stayed high. Active-matrix notebooks like the Toshiba T3400, CompuAdd Escort 450 and NEC Ultralite Versa 33C have street prices approaching $4,000, while the top-of-range IBM Thinkpad 755 will set you back even more.

And that's assuming you can find one to buy. In 1993, dealers were quoting delivery times of up to three months for some of these models. The situation has improved since then, but demand still exceeds supply, and you might still have wait several weeks to get the system you want. Because of the scarcity, there has been none of the usual aggressive discounting in this corner of the market. In fact, there have even been cases of negative discounting - street prices exceeding list prices.

Dual Scan

Inevitably, manufacturers have sought alternative display systems that would bridge the gap between the high quality of active-matrix and the low cost of passive-matrix. The most promising of these is the dual-scan system, developed by Sharp. Dual-scan notebooks started to appear in most vendors' ranges early in 1994.

A dual-scan display is essentially a passive-matrix display which is divided horizontally into two sections. It uses the same grid of transparent electrodes found in passive-matrix systems, but there are only half as many rows in each section. Because the two sections are controlled in parallel, pixels are charged for twice as long, so the display is brighter, contrast is improved and response time halved. Dual-scan screens lie roughly half way between the passive-matrix and active-matrix ends of the quality scale, and are adequate for most Windows applications.

Prices for dual-scan notebooks start at a modest $2,300 to $2,500, for basic models like the Gateway ColorBook and the Toshiba T1910CS. But higher up the market, you can pay well over $3,000, for the Compaq Contura 4/25C or the TI TravelMate 4000E, for instance. As a rough guide, a dual-scan machine costs between $600 and $1,200 less than its active-matrix equivalent.

Active Addressing

Another technology that could bridge the divide between passive- and active-matrix is the active addressing system. At the time of writing, this technology was still under development, with the first machines not expected to reach market until the end of 1994. Motif Inc, the company developing the system, claims that the quality will approach that of active-matrix, with prices falling mid-way between the levels of passive-matrix and active-matrix.

The key point about active addressing is that all the pixels are controlled simultaneously, just like in an active-matrix display, but without the need to deploy separate transistors for each of the pixels. Instead, an ASIC (application-specific integrated circuit) is used to control each row of pixels. The chip uses proprietary algorithms to calculate voltage patterns and to distribute pulses to the columns across the row.

"The immediate task facing manufacturers is to produce active-matrix displays in higher volumes and at lower costs. That should happen soon."

> *"All notebook computers use the same liquid crystal display (LCD) technology that has been around for over 20 years in watches and calculators."*

As a result, contrast ratios and clarity should be nearly as good as those of active-matrix. Response times will be around 50ms, which does not sound too good compared to the 20ms of active-matrix, but will be a lot better than the 300ms of passive-matrix. More importantly, because the production techniques are similar to those of passive-matrix, it should be possible to produce units in the quantities needed to keep up with demand.

The Future

The immediate task facing manufacturers is to produce active-matrix displays in higher volumes and at lower costs. That should happen soon. In Japan, Sharp is investing $470 million in a new factory, and Display Technologies (jointly owned by IBM and Toshiba) will triple their manufacturing capacity. In the USA, Motif recently opened the largest LCD manufacturing plant in the country, and this will eventually turn out 300,000 units per year.

As these companies solve the production problems, they will turn their attention to making larger displays with higher resolutions. At present, active-matrix notebooks only support VGA resolution, mainly with screens which are 9.5 inches across or less (measured diagonally). For today's Windows applications, users really need higher resolutions, with screens which are as large as possible, within the physical constraints of a notebook PC.

Display Technologies manufactures a 10.5-inch screen, but this has so far only found its way into products made by the company's owners, IBM and Toshiba. Panasonic has announced plans to produce larger screens, and other companies will eventually follow. That said, there will still be many users who prefer the slightly smaller screens found in today's models, as these will be lighter than the 10.5-inch version and consume slightly less power.

A more important development in active-matrix technology will be its eventual arrival on the desktop. Once manufacturers have worked out how to make a 14-inch high-resolution screen, LCDs could be on their way to replacing conventional CRT monitors. When they do, they will bring many advantages.

LCD screens will be lighter and less bulky than CRTs, they will be free of scan-line flicker and electromagnetic radiation, and they will be far more energy efficient.

Desktop LCD screens are already available from Fujitsu, NEC, Sharp and APF, but they are limited to 10.5 inches, which is too small for most PC users. But NEC has demonstrated a prototype 13-inch screen, supporting a resolution of 1280 x 1024, with 24-bit colour. Unlike other active-matrix models, it has a standard analogue interface, and can therefore be connected to an ordinary VGA port without the need for special drivers.

The Options

For the present, though, anyone shopping for a colour notebook computer has a simple set of options. If your budget is tight, you are not prepared to wait, and quality is not a factor, go out and buy one of the scores of low-cost passive-matrix models on the market. The display will be less than brilliant, but it will be good enough for most text-based applications.

If you are not prepared to sacrifice quality, consider a dual-scan model. The display will be far better than passive-matrix - certainly adequate for most Windows application - and not too much more expensive. Alternatively, wait until active addressing is available. This promises to deliver the best trade-off between quality and price.

But if you want the very highest display quality available, with sharp, bright images easily viewable from any angle, capable of supporting animation and full-motion video, don't settle for anything less than active-matrix. It is expensive, and you might have to wait a week or two to take delivery. But after a few days, you'll never want to use anything else.

PCSA

The Author

Mike Lewis is a freelance Windows specialist and writer.

How To Buy And Install CD-ROM Drives

Julian Moss offers some advice on safe and correct installation of CD-ROM drives.

CD-ROM is here to stay. Silver disks aren't just a convenient means of publishing large volumes of data, increasingly they are being used to distribute software packages and operating systems. The day is fast approaching when a CD-ROM drive will be as essential a PC peripheral as a floppy disk drive.

Choosing The Drive

The first step to adding CD-ROM to a system is choosing a drive. Performance can differ widely. The most important criterion is whether the drive is single, double, triple or quad-speed. Single speed drives have a maximum data transfer rate of 150 KB/sec - a mere 10% to 20% of that attained by an IDE hard disk.

Double speed drives can manage twice that figure. Triple speed drives achieve 450 KB/sec, and quad speed drives, which are just starting to appear, manage 600 KB/sec. Single-speed drives are, naturally, the cheapest. Apart from the speed, there's no difference between a single-speed drive and any other. You'll still be able to read the data - it will just take longer.

Seek Times

Another important factor is the average seek time. This is much worse than for hard disks, ranging from 200ms to 800ms. That's typically more than 20 times slower than an IDE drive. The reason is that not only must the head be moved to the correct track, but the speed of rotation must be varied - CD-ROMs are recorded at a constant linear velocity so, as the head moves to the outside of the disk, the speed has to be reduced. (Data on a CD, whether a CD-ROM or a music disk, starts at the middle and works outwards - the opposite of an LP record.)

Of course, the longer the seek time, the poorer the performance.

If a user is going to be making extensive use of a CD-ROM - for example, for multimedia applications - choose a fast drive. For more occasional use a slower, cheaper one may be acceptable. In many cases the performance of slower drives can be improved using caching. (The version of SmartDrive which ships with MS-DOS 6.2 will cache CD-ROMs, but make sure you load SmartDrive after the CD-ROM driver, or the cache utility won't know that there is a CD-ROM drive to cache.)

Photo CD

You'll need a drive with multi-session capability if you want to read CDs that have been recorded during separate sessions, such as Kodak Photo CDs. You may also want to ensure that the drive meets the Multimedia PC (MPC) standard. Most do, as it is such a basic standard. The latest version is MPC II. This requires a minimum of a double-speed multi-session drive with a better than 400ms access time.

> *"The loading time for small files is affected by the overhead of opening and closing them, with all the accesses to the FAT and directories that this entails."*

Interface

Once you've decided on the specification level, you must choose the type of interface. You have two options: proprietary or SCSI. However, your choice may be made for you if you need a high-specification drive: these are usually only available with a SCSI interface.

Proprietary interfaces are designed for use with a specific CD-ROM drive, and are generally easier than SCSI to set up. They are found on drives made by Panasonic, Sony, Philips and Mitsumi among others. Drives are generally supplied together with an interface - usually a small half-length ISA expansion card. If you're planning to fit a CD-ROM to an MCA machine, however, it's SCSI or nothing.

If you're upgrading a PC to full multimedia specification and fitting a sound board at the same time, it's worth noting that many sound cards include a compatible interface for popular proprietary CD-ROM drives. If you use this you'll save an expansion slot, and possibly some money as well as you may be able to buy the drive without the interface card.

SCSI, being a general purpose device interface, allows you to connect up to six other types of device, so it's a good choice for a power user's machine where advantage may be taken of these capabilities. But SCSI can be difficult to get working. *(See How to Install SCSI Devices, PCSA 65, File H0718.1.)*

> *"You'll need a drive with multi-session capability if you want to read CDs that have been recorded during separate sessions, such as Kodak Photo CDs. You may also want to ensure that the drive meets the Multimedia PC standard."*

Some vendors may offer you a CD-ROM as a kit complete with SCSI adapter. It's also possible to buy sound boards which can be upgraded to include a SCSI interface. Either of these options is an economical way to add a SCSI CD-ROM to a PC, but beware if you may later want to add other SCSI peripherals. You could find difficulty in obtaining drivers that support the bundled SCSI adapter or the sound board.

Standard Interface

Soon, you may be presented with a third choice of interface. Manufacturers are working on CD-ROM drives that can be linked to the same AT Attachment interface used by IDE hard disk drives. How this will work in practice remains to be seen. It will certainly remove the need for a separate interface card for the drive. However, system BIOS modifications will probably be needed, and you'll still have to load a device driver.

Installing The Drive

If the CD-ROM is to be fitted internally then you will need an empty 5" drive bay in the machine. The bays in most PC compatibles have holes which match the screw holes in the drive, but some machines require slide rails to be fitted. Drives aren't always supplied with these rails so, if you're going to need them, find out before you order the drive.

On others, such as some Olivetti and DEC machines, there may be no mountings at all behind the drive bay blanking plate. If extra mounting hardware is needed you'll have to obtain it from the system supplier.

To power the drive you'll need one of the free cables from the power supply. CD-ROM drives use the larger type of connector also used by hard disks and 5.25" floppy drives. If you have no spare power leads then you may be able to get a Y adapter lead with two outputs and one input, from a cable supplier.

A data cable and a separate lead for the audio output should be supplied with the drive. (Most CD-ROM drives include a headphone socket and/or external connections to allow users to listen to audio CDs while using their PC. You may wish to think about banning this practice, in order to prevent the office from turning into a music lovers' club.)

If you are installing a CD-ROM on to an existing SCSI bus then check that the ID of the drive is different from that of the host adapter or any of the other peripherals. The driver software may specify a preferred ID. If so, use it.

Terminators

Depending on where in relation to the other devices the drive is fitted, you may also need to remove a terminator from the CD-ROM drive, the host adapter, or another peripheral, to ensure that the bus is terminated only at its two ends. It's easier to remove the terminator before you install the drive. For further information about termination refer to the article in PCSA 65.

Installing The Interface

Before installing the interface card, check that the hardware resources it uses are not already used by another device. Both proprietary and SCSI interfaces will want a free IRQ line and a block of I/O addresses. A DMA channel may also be required.

Make a note of the settings if you change any of the defaults, as you will need them when you install the driver software.

If you are installing a SCSI host adapter with BIOS hard disk support, make sure that the memory region occupied by its on-board ROM is not already in use.

Now connect up the cables. Attach the power lead to the drive. When connecting the flat ribbon cable, make sure that at each end the striped edge of the cable goes to the side of the connector marked as pin 1.

The audio lead can also be connected. This lead carries audio at line voltage level. With a proprietary interface card this is brought out to a socket at the back. If you have a sound board you may want to connect it to this instead. If the drive won't be used for audio then you don't need to bother

Version	File size	DOS versions	Win 3 support?	Remarks
1.01	14,913	3.1 - 3 3	No	No ISO9660 support
2.00	18,307	3.1 - 3.3	No	High Sierra & ISO9660
2.10	19,943	3.1 - 4.0	No	Use SETVER for DOS 5.0
2.20	25,413	3.1 - 4.0	Yes	Changes to audio support
2.21	25,431	3.1 - 5.0	Yes	
2.22	25,377	3.1 - 6.0	Yes	
2.23	25,361	3.1 - 6.2	Yes	

Figure 1 - Table 1: MSCDEX Versions

CD-ROM Drives

> *"Windows for Workgroups 3.11 is worthy of special mention, because not only does it include the new SmartDrive 5.0 which can cache CD-ROMs, but it also introduces its own 32-bit file access and disk cache subsystems, which don't."*

with this cable. However, if you don't connect the cable, stick a note on the back of the PC to explain to support staff or users in the future why the audio socket doesn't work.

The Driver Software

The Driver Software will probably have an INSTALL program. If you are installing a proprietary CD-ROM drive or a SCSI drive kit complete with adapter, use this to install the software. Remember to make any changes to the default IRQ, I/O or DMA settings if you changed any jumpers when installing the interface board.

The software installation routine will make two changes to your startup files. CONFIG.SYS will be modified to include a line that looks something like this:

```
DEVICE=C:\MTMCDE.SYS /D:MSCD001 /M:20 /X
```

The name of the driver and the precise details of the switches will vary from one drive to another. However, there is some consistency between them. The /D switch generally specifies the name of an MS-DOS device by which the CD-ROM data can be accessed. An /M parameter may be supported, which specifies the number of sector buffers to be allocated - each buffer is 2 KB in size. The /X parameter in this case tells the device driver to place the buffers in extended memory.

If you are adding a CD-ROM drive to an existing SCSI bus then you will need to use ASPI or CAM drivers for all your SCSI devices, not just the CD-ROM. These may not be supplied with the drive. Non-ASPI or CAM drivers will assume that they have exclusive use of the host adapter. If you attempt to load two such device drivers then at the very least you will not be able to use the SCSI devices. More likely the PC will just crash.

If you are using ASPI, then you should have two lines in your CONFIG.SYS file like this (not excluding the drivers for other SCSI devices):

```
DEVICE=C:\ASPI4DOS.SYS /D
DEVICE=C:\MDSCD_AS.SYS /D:MSCD001
```

In this case, the first line loads the universal part of the ASPI software. This must appear above any other SCSI drivers in the file. The second line is the driver for the CD-ROM (the name will vary from one drive to another) and specifies the MS-DOS device name using the /D switch. If you are using CAM drivers then the set-up will be very similar, though the driver names will be different.

If your SCSI CD-ROM isn't shipped with either ASPI or CAM drivers then you will need to buy a separate software package. Adaptec's EZ-SCSI provides ASPI support for a number of devices, and is bundled with many of the company's adapters. Corel's CorelSCSI and CorelSCSI PRO provides ASPI support for a wider range of hardware. Future Domain's PowerSCSI! provides CAM support for Future Domain controllers, and includes an ASPI to CAM converter that enables you to use ASPI device drivers.

If you have problems getting the drive to work, check the driver version numbers. Early Adaptec drivers do not support multi-session drives.

CD-ROM Extensions

The device-specific device driver supplied with a CD-ROM drive simply provides a means to access the raw data on the CD-ROM as a DOS device. This data is generally in one of two formats: High Sierra or ISO-9660. To access the CD as an MS-DOS drive, the Microsoft CD-ROM Extensions software, MSCDEX.EXE, must be loaded. MSCDEX also provides support for audio, allowing audio tracks stored as DOS files to be played under software control.

MSCDEX was not originally included as part of DOS, so CD-ROM manufacturers generally distribute it along with their driver software. Recent versions of DOS and Windows do include the file, however. The version shipped with a CD-ROM drive is probably not the latest, so when the software installation has finished, check that you don't have a later version of MSCDEX than the one that has just been installed. If you do have a later version, on your DOS or Windows disks, use it. To find out the version number of MSCDEX.EXE, just run it by typing MSCDEX.

If you receive the message "Incorrect DOS version" when MSCDEX loads, and you don't have another copy that will work under your version of DOS, then try using SETVER to set the DOS version seen by the program to the value given in Table 1. Alternatively, you should obtain a different copy of MSCDEX from your CD-ROM drive supplier. You can also download it from CompuServe, or obtain it from Microsoft. MSCDEX is copyrighted software, however, so we cannot include it on the PCSA Utility Disk.

Command Line Options

MSCDEX has a number of command line options. The /D switch is important. This specifies the device name from which the CD-ROM data will be read, and must match the name given in the /D parameter of the CD-ROM device driver.

The /L switch lets you specify the drive letter to be used. Note that you cannot use a drive letter that has already been allocated by the time MSCDEX is run. If you have, say, a WORM drive which is assigned the drive letter D: when its device driver loads in CONFIG.SYS then you cannot use /L to override the assignment and have the CD-ROM as D: and the WORM drive as E:.

Optimising The Performance

MSCDEX has another parameter, /M, which is used to specify the number of sector buffers to be allocated. Each sector buffer takes 2 KB of memory. Your installation instructions will probably recommend that you have as large a number of buffers as you can spare, and the installation routine may set a default value of 10 or 20. However, getting the best performance out of a CD-ROM drive is a good deal more complicated than that.

CD-ROM performance is dependent on the type of access. There are two extremes: reading a single large file - equivalent to loading a large program, a graphical image or a video or audio file - or reading a number of small ones. In the second case, a lot of overhead is involved as the drive must seek constantly between the disk FAT, the directories and the file itself.

Benchmark testing reveals that buffering with the MSCDEX driver is a waste of memory unless there is no other way to cache the CD-ROM drive. This is useful because MSCDEX buffers are a memory management headache. You cannot place them in extended memory, so you must either load the driver and its buffers high, enable EMS and locate the buffers in expanded memory by specifying the /E switch, or lose a large chunk of conventional memory.

It's untrue to say MSCDEX buffers don't give any performance improvement: it's just less than you could get by other means. If the device-specific device driver supports buffering, it's much better to use this facility instead.

The time taken to load large files is limited by the data transfer rate of the drive. In fact, the additional overhead of a disk cache slows things down a little, though only by about half a per cent. SmartDrive improves performance for large files only if the cache is so large that the data is still in it the next time it is read. This won't be the case for many multimedia data files, which can be several megabytes in length.

The loading time for small files is affected by the overhead of opening and closing them, with all the accesses to the FAT and directories that this entails. The utility FASTOPEN could be a big help here: it works by holding in memory the absolute location of recently-accessed files. Unfortunately, FASTOPEN won't work with CD-ROM.

A disk cache can yield dramatic improvements, as long as it is large enough to retain a substantial amount of the CD-ROM data. However, once the CD-ROM data is flushed out of the cache by other data, file accesses will once again be held back by the need for physical seeks on the CD-ROM.

The use of separate buffering - such as that provided by the device-specific device driver - can improve performance even when a disk cache is used, because the device driver's buffers are not affected by hard disk activity. Using a 512 KB SmartDrive 5.0 cache, I found that small file access times were improved by up to 80% when 20 device driver buffers of 2 KB each were added. Since these could be placed in extended memory there was no impact on free conventional RAM.

Windows For Workgroups

Windows for Workgroups 3.11 is worthy of special mention, because not only does it include the new SmartDrive 5.0 which can cache CD-ROMs, but it also introduces its own 32-bit file access and disk cache subsystems, which don't. When you load Windows, SmartDrive reduces the size of its cache down to the second value specified on its command line: typically only 128 KB if the 32-bit cache is being used. And this is all the cache that will be available for your CD-ROM drive.

In practice, caches of 128 KB or 256 KB have little or no effect on CD-ROM access times. For users who make extensive use of CD-ROM from within Windows for Workgroups 3.11 it may be better overall to trade hard disk performance for CD-ROM speed and disable the 32-bit file access, using SmartDrive for all caching.

Conclusions

CD-ROM drives are not difficult to install. Most problems occur when installing SCSI drives, and they are usually SCSI problems.

The main headache after installation is that insufficient DOS memory may be left after the device drivers and MSCDEX have been loaded. This is almost inevitable if the PC is attached to a network. It will probably be necessary to resort to the use of QEMM or 386Max to overcome the problem.

There is no miracle cure for this under DOS. However, it is worth considering the alternative of installing the CD-ROM drives on a network server instead of individual workstations. This would make them available as network drives and avoid the need for both device drivers and MSCDEX on the workstation.

PCSA

The Author

Julian Moss is a PC support specialist and writer.

Understanding The Clipper Chip

Forget DES and RSA. If Clinton gets his way, the only method of encrypting your data will be with a device that the US Government knows how to crack.

Chances are, no one in the Clinton Administration expected that one of its major battles would be fought over something so esoteric as an encryption standard, but so it has turned out since it announced its proposal for the Clipper chip in February 1993. (It's nothing to do with the database compiler of the same name, by the way.)

Work on the Clipper chip and the algorithm, Skipjack, that it uses, started at the National Security Agency (NSA) as far back as 10 years ago. The reason was that the existing standard, DES (Data Encryption Standard), was perceived to be approaching the end of its useful life. The problem is the continuing leaps in the power of hardware: when governments can afford hardware powerful enough to crack codes by brute force, that is by trying every possible key, the standard can no longer be considered secure.

Capstone

Technically, Clipper, sometimes called Capstone after the project it's part of, is a public-key encryption system using 80-bit keys that's implemented in hardware rather than software. The chips themselves are expected to be available to manufacturers (of devices like mobile phones and modems) with NIST approval in the next couple of months.

Estimates are that the chips will cost approximately $26, adding about $100 to the price of a modem or other equipment. The NSA claims it is impossible to reverse engineer a chip to obtain its keys without destroying it.

However, there is a catch. A copy of each chip's key (and each chip has a unique key) is to be split in two, with each half deposited in escrow with a US government agency. The idea is to offer the public strong encryption, but at the same time guard the ability of law enforcement to decrypt in case of legitimate need. This means that the US government can decrypt any data that is encrypted with Clipper.

Cryptography

To understand how this could work, it's necessary to understand the nature of cryptography. The basis of encryption should be familiar to anyone who dabbled in secret codes and ciphers in their school days: you take a plaintext message and encode it using a key or password.

In secret-key cryptography, the best-known variety, the person who wants to read the message has to decode it using the same key. This causes major problems, because you now have to find a secure way of informing the message's recipient of the key. (Recursion: see recursion.)

DES, which was developed at IBM, uses secret, 56-bit keys. DES is still going strong; the NIST re-certified it for five years in 1993. Further, its life may be prolonged by multiple-encryption schemes that make it harder to crack. In the late 1970s, though, two researchers, Whitfield Diffie and Martin Hellman, came up with the idea of public-key cryptography, to get around the problems associated with trying to exchange secret keys securely.

Public-key cryptography uses pairs of keys, one public, one private. In this scheme, you encode your text using your private key, which you guard and keep secret. The person who wants to read it decodes it using your public key, which you have distributed as far and wide as you like once it has been certified as yours by witnesses who know you. There are even public-key servers where people store their keys for easy access by their correspondents.

The best-known implementation of public-key cryptography is the RSA algorithm, named after its inventors, Rivest, Shamir, and Adelman, which is used in the freeware program PGP and also in commercial programs. RSA is patented within the US, and is not supposed to be exportable - but PGP is available on BBSs and conferencing systems worldwide. Clipper is also a public-key system, and uses 80-bit keys.

Public-key cryptography has other advantages besides the one already mentioned. First of all, it offers authentication - if you can decrypt a message using someone's public key, that itself guarantees that it was encrypted using their private key. Second, if you encrypt a message using someone else's public key, only their private key can decrypt it, offering security.

To get both authentication and security, you encrypt the message using your own private key and the recipient's public key.

Speed

The downside is that public key methods are slower than secret-key cryptography, which starts to add up if you're sending long files. But public-key cryptography makes it possible to exchange those secret keys securely, and it is also inspiring a host of new applications like unforgeable digital signatures and authenticated financial transactions.

What's most likely is that we'll see a mix of methods in common use. For example, DES for large files coupled with RSA signatures and key exchange. What's certain is that encryption is one of the key technologies necessary if the promise of the data superhighways is ever to be fulfilled.

> *"Technically, Clipper, sometimes called Capstone after the project it's part of, is a public-key encryption system using 80-bit keys that's implemented in hardware rather than software."*

Family Keys

Besides the individual key, each Clipper chip has a unique serial number and an 80-bit "family key" that is common to all Clipper chips. In use, here's what is supposed to happen. The two communicating devices agree on an 80-bit session key. The message is encrypted with this key. At the same time, a law-enforcement access field is created and also sent. This includes the session key, encrypted with the unit's own key, with the sender's serial number and an authentication string appended to it. This whole string is encrypted with the Clipper family key.

A law enforcement official with a court order who wants to access this message uses the family key to decrypt the law enforcement field. This gives him the serial number of the chip and the encrypted version of the session key. The escrow agencies give him the two parts of the unit key, which he uses to obtain the session key, which gives him the ability to decrypt the message.

Standards

As of May 1994, the Clipper chip has been adopted as a voluntary US government standard. There are no public proposals to make it a mandatory standard - the White House has said that individuals, organizations, and government agencies alike are to be left free to choose other forms of encryption if they so desire. The Skipjack algorithm is classified and, although it has been examined by a panel of selected outside (although sympathetic) experts, it will not be released to the scientific community for examination.

The US government does not propose to lift export controls on strong encryption products.

It is possible that all this may change. Representative Maria Cantwell has introduced a bill into Congress that, if passed would ease the restrictions. But unless things change, and there is a lot of pressure not just from the public but from the software industry to lift restrictions, speculation is that Clipper products may be the only strong encryption products that will be readily exportable. Whether a non-US company will want to use encryption that automatically deposits a copy of the keys with the US government is another question entirely.

US citizens themselves don't like it. There is a strong movement, especially on the networks, to oppose the Clipper chip. For one thing, while today's government may be trustworthy, a whole generation of today's computer users learned the hard way in the late 1960s and early 1970s that tomorrow's government may not be. For another, while no one can criticize the Skipjack algorithm itself - because it's classified - the fear is that the economic pressure exerted by a customer the size of the government will chill the development of other, internationally competitive encryption products.

RSA Cracked?

In 1977, in a magazine article on RSA, a challenge was set. A brief message, encrypted with a 129-bit RSA key, was printed. The author challenged readers to decipher the message and offered 100 US dollars as a reward.

Last year, someone on the Internet decided to take up the challenge. He got together 800 volunteers and they shared the task between them, communicating via the internet. It took eight months for the team to crack the code, using their various computers. That's 533 years of computing time.

Various people have suggested that the success of this team spells the end for RSA. However, experts suggest that this is not the case. If you use a 512-bit or 1024-bit key for encryption (as recommended by RSA experts all along), the level of security increases by a factor of tens of thousands.

If anyone tries to tell you that RSA is useless "because it's been cracked", don't believe them.

Further Information

Electronic Frontier Foundation:
666 Pennsylvania Avenue
Washington DC, 20003, USA
Telephone +1 202 544 9237
Fax +1 202 547 5481
Email eff@eff.org
On CompuServe: GO EFFSIG.

National Institute for Standards and Technology (NIST):
Qince Orchard Road
Gaithersburg, MD, 20899, USA
Telephone +1 301 975 3058
Fax +1 301 975 2128

RSA Laboratories:
10 Twin Dolphin Drive
Redwood City, CA 94065, USA
Telephone +1 415 595 7703

PCSA

The Author

Wendy M. Grossman, a New Yorker currently living in London, England, is a freelance reporter and writer.